ARTICLE IN PRESS

Theoretical Computer Science ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Simplification of a complex signal transduction model using invariants and flow equivalent servers

Francesca Cordero [a,b], András Horváth [a], Daniele Manini [a], Lucia Napione [c,d], Massimiliano De Pierro [a], Simona Pavan [c,d], Andrea Picco [e], Andrea Veglio [c,d], Matteo Sereno [a], Federico Bussolino [c,d], Gianfranco Balbo [a,*]

[a] *Department of Computer Science, University of Torino, Torino, Italy*
[b] *Department of Clinical and Biological Sciences, University of Torino, Torino, Italy*
[c] *Institute for Cancer Research and Treatment, Candiolo (TO), Italy*
[d] *Department of Oncological Sciences, University of Torino, Torino, Italy*
[e] *European Molecular Laboratory, Research Unit of Cell Biology and Biophysics, Heidelberg, Germany*

ARTICLE INFO

Article history:
Received 3 March 2011
Received in revised form 30 May 2011
Accepted 3 June 2011
Communicated by P. Degano

Keywords:
Systems biology
Angiogenesis
Petri nets
Model simplification
Flow equivalent server

ABSTRACT

In this paper we consider the modeling of a portion of the signal transduction pathway involved in the angiogenic process. The detailed model of this process is affected by a high level of complexity due to the functional properties that are represented and the size of its state space. To overcome these problems, we suggest approaches to simplify the detailed representation that result in models with a lower computational and structural complexity, while still capturing the overall behavior of the detailed one.

The simplification process must take into account both the structural aspects and the quantitative behavior of the original model. To control the simplification from a structural point of view, we propose a set of reduction steps that maintain the invariants of the original model. To ensure the correspondence between the simplified and the original models from a quantitative point of view we use the flow equivalent method that provides a way of obtaining the parameters of the simplified model on the basis of those of the original one.

To support the proposed methodology we show that a good agreement exists among the temporal evolutions of the relevant biological products in the simplified and detailed model evaluated with a large set of input parameters.

1. Introduction

Formal modeling is a central theme in systems biology where mathematical reasoning and simulation can play an important role. In general, biological models are affected by a high level of complexity due to the organization and functional properties of the systems that are considered. The interaction of qualitative and quantitative analysis is necessary to check a model for consistency and correctness. Following this idea, Heiner et al. proposed in [18] a methodology to develop and analyze large biological models in a step-wise manner. This approach consists of four steps focusing on: (1) readability, (2) executability (animation techniques), (3) validation techniques and (4) analysis techniques. During the modeling and analysis processes these stages could represent a supervised way to increase the confidence on the results provided by the model.

* Corresponding author. Tel.: +39 0116706740.
E-mail address: balbo@di.unito.it (G. Balbo).

0304-3975/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2011.06.013

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

In our opinion, between the exploration of executability of the model and the validation of the model's integrity (second and third steps, respectively) it is necessary in some cases to apply a simplification step. This step becomes mandatory considering two main obstacles in the analysis of biological systems: (i) the number of states of a system may grow exponentially with the number of compounds involved in the representation, and (ii) the large number of kinetic parameters that are needed, many of which are unknown.

For these reasons, simplification is an important issue of the model analysis process that, starting from the construction of the original detailed description, attempts to identify only those variables that are crucial to the dynamics of the system in order to obtain a representation suitable for making predictions, generating new hypotheses, and suggesting the design of novel biological experiments (see [8]).

As reported in [16] there are two main classes of reduction techniques that can be used: *exact*, based on the idea of grouping together states with similar characteristics so that the behavior of the whole model remains un-affected by this state space modification, and *approximate*, based on a deep understanding of the features of the system and yielding only approximate results. In each type of approach the conservation of stoichiometric and kinetics aspects of the biochemical system must be ensured. From a biochemical point of view, several proposals have been presented [39,34] to describe different methods for the conservation analysis based on the use of the stoichiometry matrix in order to detect the presence of dependences.

These general techniques may also be interpreted as two-phase procedures. The first phase concerns a biological analysis of the system: the compounds that have a significant effect on the behavior of the system are necessary while compounds and parameters that have a weak influence may be eliminated leading to a reduced model [40]. The second phase relies on specific features of the formalism used to represent the biological system and reduces the size of the state space of the model by identifying structures of the state space (e.g., symmetry conditions) that allow the analyst to group together states with similar or identical behaviors [6,5].

In this paper we report our experience of modeling signal transduction pathways for the angiogenic phenomenon where we devised a complexity reduction methodology applied to the detailed representation of the dynamics of the system. We model this phenomenon using Petri Nets (PNs) [33] (and in some cases their variant called Stochastic Petri Nets (SPNs) [30,27,4]).

We propose a simplification process which takes into account both the biochemical and the formalism views.

Similarly to what has been done in other proposals, we use structural properties of the PN model to assist the quantitative analysis. In particular, our simplification process is based on the identification of basic building blocks (groups of biological reactions) that can be replaced in the PN representation with equivalent "macro" reactions which preserve the overall behavior of the model. This reduction process is guided by qualitative properties whereas the accuracy of the substitutions is validated by comparing the quantitative characteristics of the detailed and simplified models. In deriving information related to the qualitative analysis, an important role is played by the so-called net's *invariants* [33,28].

The quantitative validation is performed by translating the PN models into the corresponding sets of coupled, first order Ordinary Differential Equations (ODEs).

Obviously, the specifications of the reduced models depend on the dynamics of the original one. To perform this transformation, we apply the concept of *flow equivalent server* [7,10,36], investigating the conditions which ensure that the behaviors of the simplified nets are similar to that of the detailed one.

The discussion contained in this paper is focused on the robustness of the approach and is based on numerical and analytical results that have been derived without referring to parameters coming from wet-lab experiments. Indeed, even though the chronology of the biochemical events is properly defined, the rates and concentration values are not yet available, consequently our knowledge about the enzyme kinetic mechanism and the quantitative aspects that dominate all reactions is limited. Within this context, our work must be considered as a preliminary step toward the application of the proposed methodology to realistic and complex models. In particular, studying the reliability of the method performed with respect to a parameter space characterized by wide ranges of the individual parameters, allows us to conclude with the belief that our methodology is a promising approach in the direction of attacking a problem that is crucial in systems biology.

The paper, which is an extended version of [29] that uses the flow equivalent method for computing the parameters of the reduced networks, is organized as follows. Section 2 provides an overview of PNs and SPNs and of their use in biochemical systems for qualitative and quantitative analysis. Section 3 describes the angiogenic case study and presents the approach we adopted to build the SPNs. Section 4 shows the simplification processes we followed. Section 5 discusses the validity of the resulting models from a structural point of view. In Section 6, the attention is focused on the quantitative issues of the simplification process. The accuracy and the mathematical robustness of the simplified model are presented in Section 7. We conclude with a discussion and an outlook of future work in Section 8.

2. Modeling formalism and solution techniques

The descriptions commonly used in biology, where the relations among reagents are expressed either by biochemical reactions, or by interactions of genes, are easy to transform into PNs in which places correspond to genes/proteins/-compounds (substrates) and transitions to interactions. PNs have been first proposed for the representation of biological pathways by Reddy et al. [32]. Since their introduction, many other researchers proposed PN models of biological systems [20] with the aim of using their representations to obtain quantitative information about the behavior of these

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

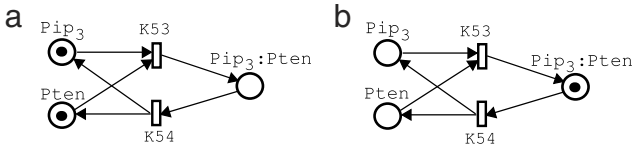

Fig. 1. PN representation of reactions $Pip_3 + Pten \underset{k_{54}}{\overset{k_{53}}{\rightleftarrows}} Pip_3{:}Pten$.

systems, mostly via simulation [21,14]. However, as shown by Heiner et al. [18,19], PNs can be used for deeper studies of biological systems where both qualitative and quantitative results are derived for assessing general behavioral patterns and specific concentration or timing measures. Indeed, the interaction of qualitative and quantitative analysis is necessary to check a model for consistency and correctness as we will show in the rest of this paper.

2.1. Petri net representation for biochemical interactions

PNs are a graphical language for the formal description of distributed systems with concurrency and synchronization. PNs are bipartite graphs with two types of nodes, namely places and transitions, connected by directed arcs. The state of the system is given by the distribution of tokens over the places of the net and is called *marking*. The dynamics of the model (starting from an initial marking) is captured by state changes due to firing of transitions and by the consequent movement of tokens over the places. Formally, a PN is characterized in the following manner.

Definition 1 (*PN-syntax*)**.** A Petri Net is a tuple $(P, T, W, \mathbf{m}_0)$, where

- P is a finite set of *places*;
- T is a finite set of *transitions*;
- P and T are such that $P \cap T = \emptyset$;
- $W : (P \times T) \cup (T \times P) \to \mathbb{N}$ defines the arcs of the net and assigns to each of them a multiplicity, thus representing the flow relations of the net; a particular element $W(p, t)$ (or $W(t, p)$) gives the number of tokens that the firing of transition t takes away from (or puts into) place p;
- $\mathbf{m}_0$ is the *initial* marking which associates with each place a number of *tokens*.

Given the structure of the net, the dynamics is described by the evolution of its marking that is governed by transition firings which remove tokens from input places and put tokens in output places. Identifying with ${}^\bullet t$ the set of input places of transition t (${}^\bullet t = \{p : w(p, t) > 0\}$), we say that transition t is *enabled* if all of its input places are marked with at least as many tokens as the multiplicities of the corresponding arcs. More formally, we can say that t is *enabled* in marking $\mathbf{m}$ *iff* $\{\forall p \in {}^\bullet t : m(p) \geq w(p, t)\}$. The firing of transition t, enabled in marking $\mathbf{m}$, yields a new marking $\mathbf{m}'$ such that $\forall p \in P : m'(p) = m(p) - w(p, t) + w(t, p)$.

The set of all the markings that the net can reach, starting from $\mathbf{m}_0$, is called the reachability set of the net and denoted with $RS(\mathbf{m}_0)$.

As already observed, when PNs are used in systems biology places represent biochemical entities (enzymes, compounds, etc.) and transitions correspond to their interactions [32]. We assume that the tokens in the places represent the number of molecules of the corresponding entities. The biological system we consider is described by biochemical reactions similar to those reported in Fig. 1(a) where we show the PN representation scheme we adopted to model all reactions of this type. Fig. 1(b) represents the state evolution of this elementary sub-net, due to firing of transition K_{53} (occurrence of the forward reaction with rate $k53$).

2.2. Analysis techniques based on structural properties

The PN graph representation is the basis for the computation of several functional properties of the model, that are valid independently of its initial marking: such properties are, for instance, the boundedness, the occurrence of structural deadlocks and traps, and the potential existence of *home* states, i.e., the possibility for the net of returning infinitely often to its initial state [33,28,4]. In deriving such kinds of information, an important role is played by certain properties of the net which are called *invariants*. There exist two kinds of invariants: place invariants (P-invariants) and transition invariants (T-invariants) [33,28].

Definition 2 (*Semiflows*)**.** Given a Petri Net, let

- $\mathbf{C}$ be the *Incidence Matrix* whose generic element $c_{pt} = W(t, p) - W(p, t)$ describes the effect of the firing of transition t on the number of tokens in place p;
- $\mathbf{x} \in \mathbb{Z}^{|P|}$ be a *place* vector;
- $\mathbf{y} \in \mathbb{Z}^{|T|}$ be a *transition* vector;
- A *P-semiflow* be a place vector $\mathbf{x}$ that is an integer and non-negative solution of the matrix equation $\mathbf{xC} = \mathbf{0}$;
- A *T-semiflow* be a transition vector $\mathbf{y}$ that is an integer and non-negative solution of the matrix equation $\mathbf{Cy} = \mathbf{0}$.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

The support of a semiflow **h**, denoted with *supp*(**h**), is the set of nodes corresponding to the non-zero entries of **h**.
A *P-invariant* is a weighted sum of tokens contained in a subset of places of the net that remains constant through the entire evolution of the model, starting from an initial marking. Each *P*-semiflow **x** allows the computation of a corresponding *P*-invariant using *supp*(**x**) as the subset of places and the non-zero entries of **x** as weights. All the *P*-semiflows of a PN can be expressed as linear combination of a set of *minimal P*-semiflows that are thus the basis for the analysis [25,28]. When the net is covered by *P*-semiflows (i.e., all the places of the net belong to one *P*-semiflow, at least) the state space of the net is finite and the markings of the places are bounded. The interpretation of a (minimal) *P*-invariant in a biological context, where tokens represent compounds, enzymes etc., is relatively simple: the places of *supp*(**x**) represent the portion of the PN where a given kind of correlated matter is preserved.

Slightly more complex is the definition of a *T*-invariant that corresponds to a set of transitions whose firing may bring the net back to its initial marking. *T*-invariants are interpretations of *T*-semiflows **y**. Recall that [33], given a marking **m**, the firing of transition t_i yields a new marking $\mathbf{m}'$ which results from the addition of two vectors: the marking **m** and the column $\mathbf{c_i}$ corresponding to transition t_i in the incidence matrix **C**. The effect of the firing of all the transitions of a *T*-semiflow is given by the weighted sum $\sum_{t_i \in T} y_i * \mathbf{c_i}$. It follows from its definition that a *T*-invariant corresponds to the situation in which the previous sum yields a null vector and implies that the total modification of the marking is null after the firings of all the transitions involved by the semiflow (i.e., identified by the support of *T*-semiflow **y**). In other words, the non-zero components of a *T*-semiflow can be interpreted as the number of times the corresponding transitions must appear in a sequence of firings that brings the net back into a given marking.

As in the case of *P*-semiflows, all the *T*-semiflows of a PN can be expressed as linear combination of minimal *T*-semiflows [28]. *T*-semiflows represent invariant laws derived from the structure of the net (the PN graph), but knowledge of them alone is not sufficient to state the effective enabling of the transitions they involve and further considerations are necessary to study, for instance, liveness properties or cyclic behaviors of the system. In a biological context, *T*-semiflows can be interpreted as *elementary modes* as discussed at some length in [15].

2.3. Quantitative temporal analysis

One main objective in systems biology is to model and analyze temporal dynamics of the phenomenon under study. It is natural hence to apply an extension of PN that allows the introduction of temporal specifications in the model. The most common timed extension of PN is SPN in which exponentially distributed random delays (interpreted as durations of certain activities) are associated with the firings of the transitions. SPNs were originally defined in [27,30]

Definition 3 (*SPN-syntax*)**.** A Stochastic Petri Net is a pair (PN, K), where

- PN is a Petri net;
- K is a function which allows the definition of the stochastic component of an SPN model, mapping transitions into real positive functions of the SPN marking (rates of the corresponding negative exponential distributions).

Given this definition, for any transition t it is necessary to specify a function $K(t, \mathbf{m})$. In the case of marking independency, the simpler notation k_i is normally used to indicate $K(t_i)$, for any transition $t_i \in T$. When transition t_i is enabled in marking **m**, evaluating the function $K(t_i, \mathbf{m})$ (or k_i) provides the *rate* of transition t_i in marking **m**, i.e., the parameter of the negative exponential distribution that characterizes the firing time of this transition in this marking. Assuming that the *atomic* firing (characteristic of the un-timed model) is preserved in the SPN specification (i.e., once enabled, the transition waits for a random delay before actually firing, then removing tokens from its input places and adding tokens to its output places in zero time) and given that the distribution of the firing times have infinite supports, SPNs are qualitatively equivalent to PNs, meaning that the RSs of the two models are identical and that for their structural analysis it is sufficient to disregard their time specifications. The temporal stochastic behavior of an SPN is isomorphic to that of a continuous time Markov chain (CTMC) which can be built automatically from the description of the SPN. In particular, bounded SPNs can be shown to be isomorphic to finite CTMCs. This *stochastic approach* based on SPN adopts a discrete view of the quantity of the entities and sees their temporal behavior as a random process governed by the so-called Chapman–Kolmogorov differential equations [11], and corresponding to the behavior of the biological system described by the Master Chemical Equations [13].

Starting from the information contained in the SPN model, it is also possible to adopt a deterministic approach in which the temporal behavior of the quantity of the entities contained in the different places is a completely predictable process. In the following we give a brief description of this approach; for a detailed discussion, the interested reader is referred to [17].

The deterministic approach translates the interactions into ODEs with one equation per place. When modeling metabolic pathways, the most common way to translate the reactions into a set of ODEs is provided by the law of generalized mass action (GMA) [41]. By GMA the system of ODEs describing the model is of the form

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^{N_i} k_{ij} \prod_{h=1}^{E} X_h(t)^{g_{ijh}} \quad (i = 1, \ldots, E) \tag{1}$$

where E is the number of interacting entities and $X_i(t)$ represents the amount of the ith entity at time t. Furthermore, N_i is the number of reactions in which the ith entity is involved, the parameters k_{ij} are rate constants describing the speeds

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

of these reactions and the parameters g_{ijh} are the so-called *kinetics orders* which depend on the stoichiometry and on the mechanism of the reactions. As it is possible to deduce from (1), the behavior of the entities involved in the system depends both on the network topology and on the reaction rates.

Referring again to the reactions considered in Fig. 1, the corresponding ODEs are

$$\begin{aligned}
\frac{dX_{Pip_3}(t)}{dt} &= -k_{53}X_{Pip_3}(t)X_{Pten}(t) + k_{54}X_{Pip_3:Pten}(t),\\
\frac{dX_{Pten}(t)}{dt} &= -k_{53}X_{Pip_3}(t)X_{Pten}(t) + k_{54}X_{Pip_3:Pten}(t),\\
\frac{dX_{Pip_3:Pten}(t)}{dt} &= k_{53}X_{Pip_3}(t)X_{Pten}(t) - k_{54}X_{Pip_3:Pten}(t).
\end{aligned}$$

The ODEs can be obtained automatically from the SPN representation[1] and having also the information about the initial amount of the different entities, numerical integration of the ODEs is applied to calculate the quantities at a given time instant.

As is shown in [22], when the number of tokens increases the quantitative behavior obtained applying the stochastic approach tends to that obtained from the ODEs. Having the SPN description of the biological phenomenon, the construction of the corresponding CTMC or a set of ODEs can be done automatically. The choice of using one of the two approaches (stochastic or deterministic) for studying the behavior of the system is thus left to the analyst who decides on the basis of the objectives of his/her study. In this paper we mainly use the deterministic approach because it allows for faster and simpler evaluation of the simplification process we propose for the SPN representation of our system.

3. A Petri Nets based approach applied to signal transduction pathways for the angiogenic process

One main objective in systems biology is to model and analyze the temporal dynamics of the phenomenon under study. By using SPNs as the formalism for the construction of the model, the analysis is performed in two phases: the first provides qualitative information about the structure of the model and the second investigates its quantitative behavior including the computation of statistical indexes that describe the dynamics of the system. In this paper we use this approach to study an angiogenic signal transduction system described in Section 3.1 and modeled in Section 3.2.

3.1. Biological case study definition

The example used in this paper describes the intracellular signal transduction events induced by Vascular Endothelial Growth Factor (*VEGF*) in the context of the angiogenic process. Angiogenesis is a complex phenomenon that proceeds from a molecular level through specific cellular mechanisms to macroscopic events defined as the formation of new vessels from preexisting ones. At the molecular level this phenomenon involves the activities of many growth factors and relative receptors which trigger several signaling pathways resulting in different cellular responses. *VEGF* family proteins are widely regarded as the most important growth factors involved in angiogenesis. Among the *VEGF* family proteins, *VEGF-A* has been the subject of many investigations and is well recognized as the major angiogenic factor. *VEGF* receptor-2 (*KDR* in humans) is the primary mediator of *VEGF-A*-induced cellular responses, including cell proliferation, survival, and migration [31]. Although the core components of the main *KDR*-induced pathways have been identified, further research is needed to better elucidate the *KDR*-signaling network. Indeed, a strong body of evidence indicates the existence of common adaptor/effector proteins involved in the survival and proliferation pathways induced by *VEGF-A*/*KDR* axis, pointing out the difficulty of isolating a specific pathway and suggesting the presence of common nodes which contribute to create an intricate signaling network. In particular, the phosphorylated active receptor, indicated as *KDR**, catalyzes phosphorylation of several intracellular substrates including the adaptor protein *Gab*1 [23,9], which is shared by both proliferative and survival signaling. The main pathway through which *VEGF-A* induces cell proliferation involves the activation of Plc_γ [37]. Activation of Plc_γ promotes phosphatidylinositol 4,5-bisphosphate (Pip_2) hydrolysis giving rise to 1,2-diacylglycerol (*Dag*). *VEGF-A*-induced cell survival is dependent on the activity of *Pi3k* [12]. The activated *Pi3k* phosphorylates Pip_2 generating phosphatidylinositol-3,4,5-triphosphate (Pip_3). This recruits *Akt* to the membrane where it is activated through phosphorylation. Activated *Akt* induces cell survival. Taking into account these notions, we wrote a system of biochemical reactions based on the available biological information together with further supposed mechanisms which could contribute to underline the presence of key molecular nodes in the context of *VEGF-A*-induced proliferation and survival pathways.

[1] This automatic transformation is implemented in several analysis tools such as GreatSPN tool [3] which has been used to support many of the analyses discussed in this paper.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

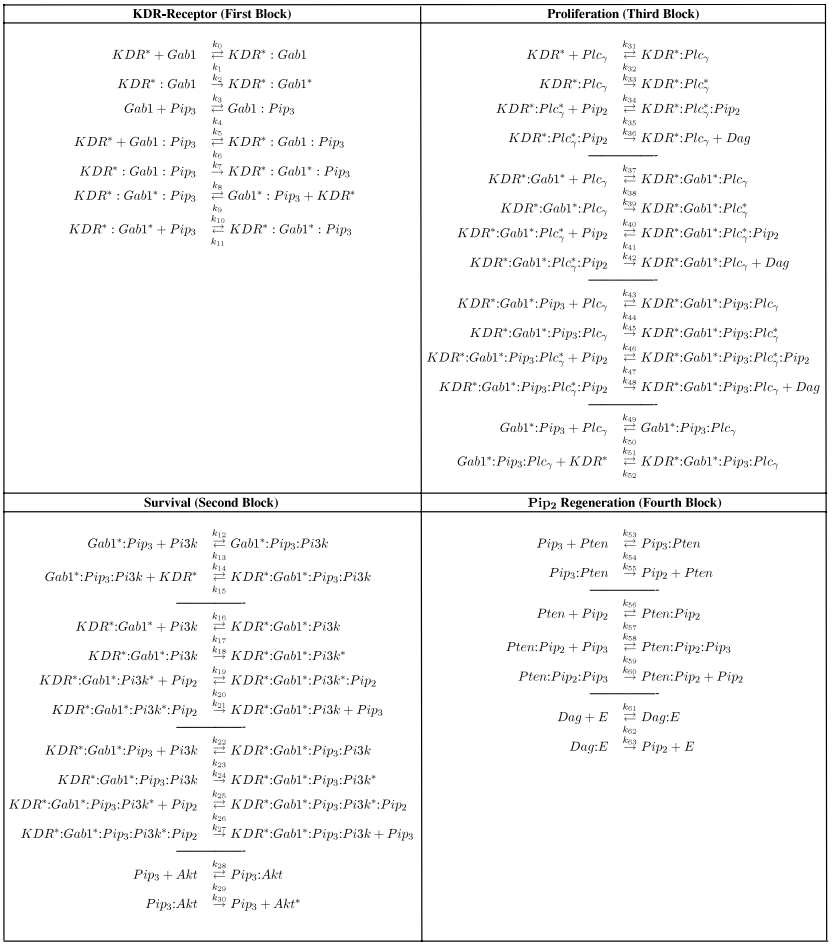

Fig. 2. Reactions of the detailed model.

3.2. Model construction

In this section, we discuss the approach we followed to represent the signal transduction cascade using an SPN. Consider the detailed biological model depicted in Fig. 2 and comprising 63 reactions. Note that in the following we denote a reaction rate with k_i, whereas the corresponding transition in SPNs is named with K_i and its firing rate is equal to $k_i(\mathbf{m})$ where $\mathbf{m}$ is the marking of the input places of the transition and the tokens denote the number of molecules.

These reactions describe *KDR*-proximal signaling events in the context of the survival and proliferation signal modules induced by receptor activation. In particular, reactions are organized into four blocks. The First Block (of Fig. 3) represents the earliest signaling events which include *KDR** (we use the star to denote that proteins are active), *Gab*1, and Pip_3. The Second Block includes the reactions describing the survival pathway triggered by the *Pi*3*k*/*Akt* axis. The Third Block represents the proliferation pathway involving Plc_γ activation. Finally, the Fourth Block concerns the regeneration of Pip_2, a common substrate for the two signal modules that we are considering. In this block Pip_2 recovery was considered to result from the contribution of *Pten*-dependent dephosphorylation of Pip_3 in combination with *Dag* catabolism (here recapitulated in the general enzyme *E*). Using the reaction representations outlined in Section 2.1 and the GreatSPN tool [3] the SPN model of the signal transduction process was built as illustrated in Fig. 3. Exploiting the block organization and the structure of the model we analyzed the biochemical reactions in order to identify possible pathways and sub-pathways that describe embedded behaviors of the complete model. We denoted the reactions by means of their kinetic constants. In the model, *Akt* and *Dag* have been considered as the end points of the survival and proliferation pathways, respectively. Taking into account these end points in combination with the notion that *Akt* activation is strictly Pip_3-dependent, we examined the signal transduction cascade focusing our attention mainly on the reactions that lead to the production of Pip_3 (i.e., K_{21} and K_{27}) and *Dag* (i.e., K_{36}, K_{42}, and K_{48}).

This analysis (supported also by a careful drawing of the SPN) allowed us to recognize different sub-pathways that lead to the survival or proliferation effects. In the context of the survival signal module, we identified three sub-pathways, that

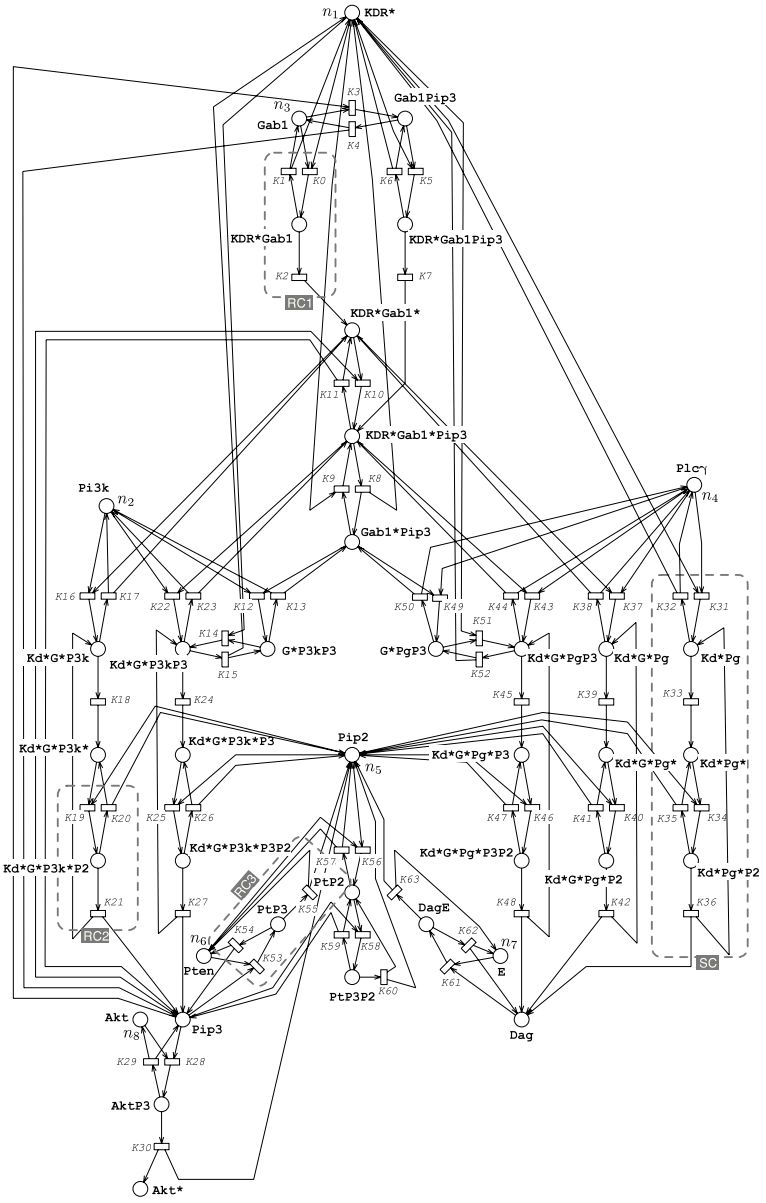

Fig. 3. SPN representing the detailed model. Compound abbreviations: $KDR^* \equiv Kd^*$, $Gab1 \equiv G$, $Pi3k \equiv P3k$, $Plc_\gamma \equiv Pg$, $Pip_3 \equiv P3$, $Pip_2 \equiv P2$, $Pten \equiv Pt$. (For the sake of simplicity, abbreviations are used only for compounds bound with the enzyme).

lead to a survival effect starting from: KDR^*:$Gab1^*$, KDR^*:$Gab1^*$:Pip_3 and $Gab1^*$:Pip_3. Two of them are characterized by the presence of a distinguishing complex, KDR^*:$Gab1^*$ or KDR^*:$Gab1^*$:Pip_3, belonging to the First Block. Moreover, $Gab1^*$:Pip_3 also contributes to the formation of KDR^*:$Gab1^*$:$Pi3k$:Pip_3 complex already involved in one of the identified sub-pathways. Summarizing there are three sub-pathways that lead to a survival effect starting from: KDR^*:$Gab1^*$, KDR^*:$Gab1^*$:Pip_3 and $Gab1^*$:Pip_3.

Turning our attention to the proliferation module, we identified four different sub-pathways that are distinguished by the compounds belonging to the First Block, i.e.: KDR^*, KDR^*:$Gab1^*$, KDR^*:$Gab1^*$:Pip_3 or $Gab1^*$:Pip_3. Notice that the distinguishing elements of the detected sub-pathways are the same within the survival and proliferation modules, with the exception of the compound KDR^*.

Referring again to the SPN of Fig. 3, we can notice that the time evolution of this SPN is intuitively portrayed by a top-down view. On the top is depicted the place KDR^* that represents the starting point of the signal cascade induced by its ligand. All the sub-pathways that characterize the proliferation and survival pathways start from the KDR^* cascade. The places describing Dag and Pip_3 are aligned at the bottom of the net.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

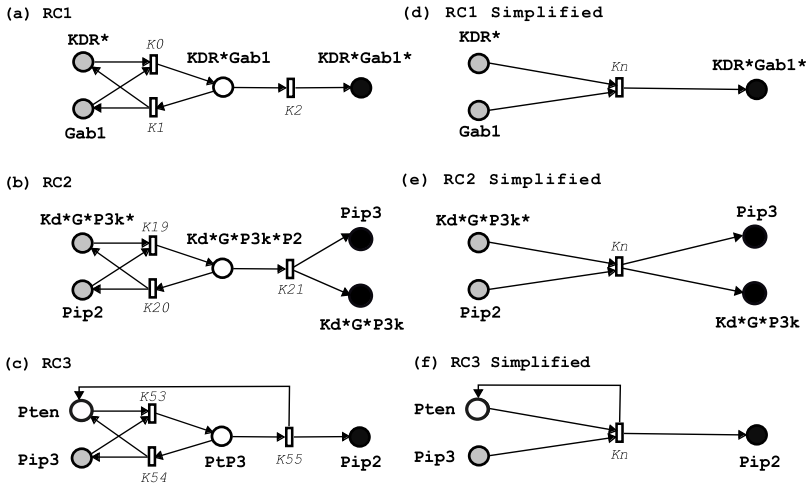

Fig. 4. Subtypes of the component RC and its simplified version. Note that K_n is equivalent to K_0, K_{57}, and K_{62} for (d), (e), and (f) respectively in Fig. 5.

As reported in [29] the resulting SPN is covered by *P*-invariants (see Table A.2 in Appendix).

4. Model simplification

The molecular mechanisms involved in the *VEGF-A* angiogenic functions need to be characterized in fine details in order to better understand the flow of information among cell proliferation, survival, and migration. Experimental evidence demonstrates the existence of common adaptor/effector proteins in both the pathways analyzed in this paper. These shared compounds form an intricate network that is worthy of thorough investigation. The SPN model representation (Fig. 3) takes into account the presence of these competitions and offers a tool for investigating the temporal dynamics of two relevant products in the context of *VEGF-A*-induced proliferation and survival pathways, such as *Dag* and Pip_3 respectively.

The analysis of the model requires knowledge of the transition rates (that in the biological context correspond to the reaction rates) and of the initial marking (the initial quantities of the chemical compounds).

The quantitative analysis of this SPN is hampered by the size of its state space that can grow very large and suffers from numerical problems that are quite common in models of this type (e.g. stiffness).

To mitigate these problems we propose a model simplification process that aims at the reduction of the structural complexities of the model without losing its capability of representing the overall behavior of the considered biological phenomenon. This approach identifies the presence of sub-models (sub-nets) with similar structures, that we call *components*, and replace them with simpler constructs without heavy modification of the overall properties of the model. In particular, the simplification process eliminates the intermediate proteins complexes (e.g. *Kd**:*Pg* and *Kd**:*Pg**:*P*2) and it allows an analysis focused on the common adaptor and effector proteins of the cascade (e.g. Plc_γ and *KDR**:*Gab*1 respectively). The resulting model is characterized by a simpler structure that can better support the analysis of the system.

In Section 4.1, we perform the simplification process based on the replacement of a small component that is replicated many times in the detailed model. We show its structure, describe the methodology adopted to simplify it and present the resulting overall simplified model. In Section 4.2, we propose a more drastic simplification process which is based on the replacement of a larger component. In Section 5, we verify that the simplification processes preserve the structural properties of the original model.

4.1. Simplification process: first reduction component

We identified a small component, which we call Reduction Component (RC), that is replicated many times in the model. All the occurrences of RC share a similar basic structure with small variants that suggest to group them into three different classes which we discuss individually in the following, denoting the *input* places by gray places and the *output* places by black places.

The first type has the structure depicted in Fig. 4(a) and is highlighted in Fig. 3 as RC1. The behavior of this sub-model in isolation can be summarized as follows. Place *KDR**:*Gab*1* is absorbing, i.e., the tokens present in place *KDR**:*Gab*1* cannot move away from this place. Tokens positioned in places *KDR** and *Gab*1 have instead the effect of starting possible evolutions of the sub-net which result into tokens eventually appearing in place *KDR**:*Gab*1*. This interpretation is supported by the structural analysis of the sub-model which shows the presence of two minimal *P*-semiflows with identical supports, except for the presence of *KDR** in the first and of *Gab*1 in the second.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

The second type (denoted as RC2 in Fig. 3) differs from RC1 since it has two output places as reported in Fig. 4(b). Besides this "double" output, RC2 behaves exactly as RC1. In this case, there are four P-semiflows that take into account the flows of tokens starting from each input place to each output place.

The third type, denoted as RC3 in Fig. 3, is reported in Fig. 4(c). In this sub-model the only input place is Pip_3 and the only output place is Pip_2. The difference with respect to RC1 is the presence of the arc from transition K_{55} to place *Pten* that behaves as a feedback, moving the tokens produced by the *output* transition K_{55} to the place *Pten*. This difference is reflected in the P-semiflows of the sub-model, one of which covers places *Pten* and $Pt{:}P3$ while the other comprises places Pip_3, $Pt{:}P3$, and Pip_2. This means that the net effect of this sub-model is that of transforming Pip_3 into Pip_2 leaving to *Pten* the role of an enabling component.

Even if the components RC1, RC2 and RC3 are not identical, we can simplify them in a similar manner taking into account their structural differences. In all types of the RC sub-net tokens flow from *input* places toward *output* places and this can be captured by a single transition.

The resulting simplified sub-nets are depicted in Fig. 4(d)(e)(f) where transitions K_n play the role of moving the tokens from the *input* places toward the *output* places.

The feedback arc in RC3 (from transition K_{55} to place *Pten*) is included in the simplified sub-net as well, where we have an arc from transition K_n to the appropriate *input* place (see Fig. 4(e)).

From a biological point of view, the RC1 corresponds to the first group of reactions after the activation of KDR^*. They are characterized by reversible reactions indicated with transitions K_0 and K_1 and one-way reactions represented with transition K_2:

$$KDR^* + Gab1 \underset{k_1}{\overset{k_0}{\rightleftarrows}} KDR^*{:}Gab1 \xrightarrow{k_2} KDR^*{:}Gab1^*.$$

In the original model this structure describes a two-step procedure composed by (i) the interaction between the receptor KDR^* and an inactive protein, and (ii) the protein activation. The simplification procedure leads to the substitution of the three reactions with a single step (a single reaction from substrates directly into products) that appears as follows:

$$KDR^* + Gab1 \xrightarrow{k_n} KDR^*{:}Gab1^*.$$

The RC2 structure characterizes the reactions that produced Pip_3 and *Dag*. RC2 is similar to RC1 except for the one-way reaction that has two output places.

$$Kd^*{:}G^*{:}P3k^* + Pip_2 \underset{k_{20}}{\overset{k_{19}}{\rightleftarrows}} Kd^*{:}G^*{:}P3k^*{:}P2 \xrightarrow{k_{21}} Kd^*{:}G^*{:}P3k + Pip_3.$$

The reduction mechanism is the same of RC1:

$$Kd^*{:}G^*{:}P3k^* + Pip_2 \xrightarrow{k_n} Kd^*{:}G^*{:}P3k + Pip_3.$$

Differently from the previous cases, in the RC3 structure the two-step procedure is enriched with a feedback arc characterizing reactions in which the enzyme after the catalysis returns in the free form. The detailed reactions are

$$Pten + Pip_3 \underset{k_{54}}{\overset{k_{53}}{\rightleftarrows}} Pt{:}P3 \xrightarrow{k_{55}} Pten + Pip_2$$

that are transformed in a single step:

$$Pten + Pip_3 \xrightarrow{k_n} Pten + Pip_2.$$

The simplification process identifies all the components of the same type and replaces them with their corresponding reduced structures. The SPN obtained after the application of this simplification process is depicted in Fig. 5.

4.2. Simplification process: second reduction component

Also a larger repetitive sub-net can be identified in the model. We will refer to this as Second Component (SC), which is highlighted on the right side of Fig. 3 and shown in Fig. 6(a) in isolation. The sub-nets characterized by this structure correspond to a specific biochemical cascade:

- the binding between an enzyme and the first substrate (transitions K_{31} and K_{32});
- the enzyme activation (transition K_{33});
- the recruitment of the second substrate (transitions K_{34} and K_{35});
- the production of the molecules representing the pathway end point and the enzyme deactivation (transition K_{36}).

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

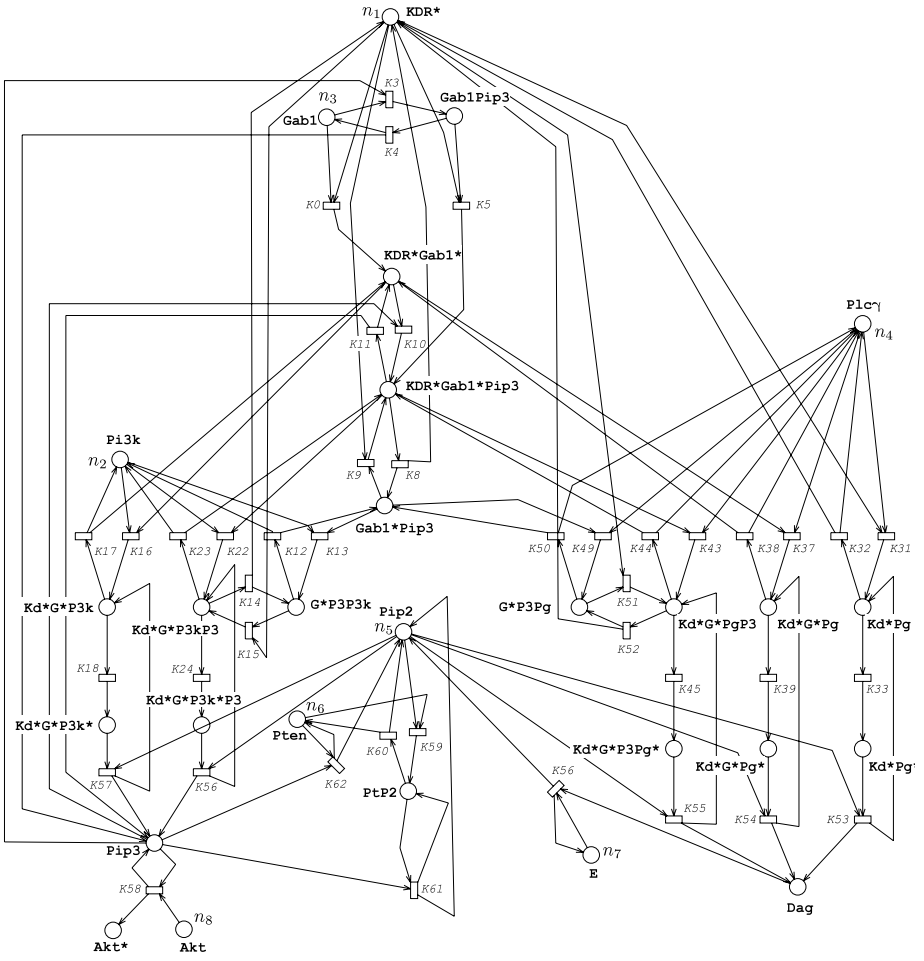

Fig. 5. SPN obtained after the first process of simplification.

All SC sub-nets present in the detailed SPN correspond to an enzymatic reaction group of the form

$$Plc_\gamma + KDR^* \underset{k_{32}}{\overset{k_{31}}{\rightleftarrows}} Kd^*{:}Pg \xrightarrow{k_{33}} Kd^*{:}Pg^*$$

$$Kd^*{:}Pg^* + Pip_2 \underset{k_{35}}{\overset{k_{34}}{\rightleftarrows}} Kd^*{:}Pg^*{:}P2$$

$$Kd^*{:}Pg^*{:}P2 \xrightarrow{k_{36}} Kd^*{:}Pg + Dag$$

which is then reduced to the following single step

$$Plc_\gamma + KDR^* + Pip_2 \xrightarrow{k_n} Plc_\gamma + KDR^* + Dag.$$

The second simplification approach that we propose aims at reducing the entire SC component represented in Fig. 6(a) with a single transition appropriately connected with the rest of the model. The same flow arguments used to justify the reduction step of the first simplification process allow the identification of Pip_2 as the only *input* place (gray place in Fig. 6), while again only place *Dag* can be considered as an *output* place (black place in Fig. 6). Also in this case, these equivalence arguments are supported by the structural analysis of the sub-model that identifies three P-semiflows. The first that covers places Pip_2, $Kd^*{:}Pg^*{:}P2$, and *Dag*, singles out the flow from Pip_2 to *Dag*. The other two which are identical, except for their first component ($\{Plc_\gamma, Kd^*{:}Pg, Kd^*{:}Pg^*, Kd^*{:}Pg^*{:}P2\}$ and $\{KDR^*, Kd^*{:}Pg, Kd^*{:}Pg^*, Kd^*{:}Pg^*{:}P2\}$) highlight the fact that places KDR^* and Plc_γ play the role of enabling compounds.

With this in mind, the six transitions of the detailed sub-model are substituted with a single transition of the reduced model which is K_n as shown in Fig. 6(b) and all the intermediate protein complex are eliminated.

By performing the reduction of both the SC type sub-nets and the remaining RC type ones, we obtain the net depicted in Fig. 7.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

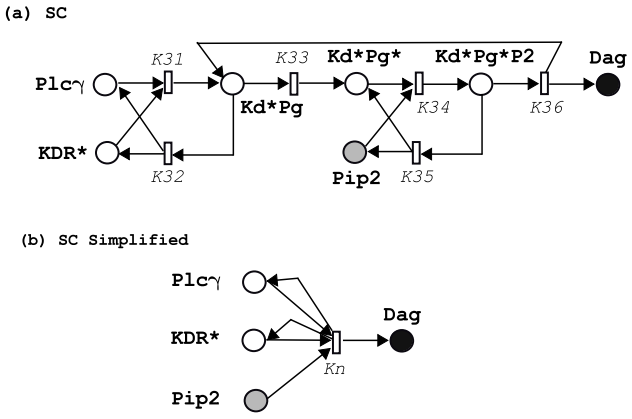

Fig. 6. SC sub-net and its reduced counterpart. Note that transition K_n corresponds to K_{15} in Fig. 7.

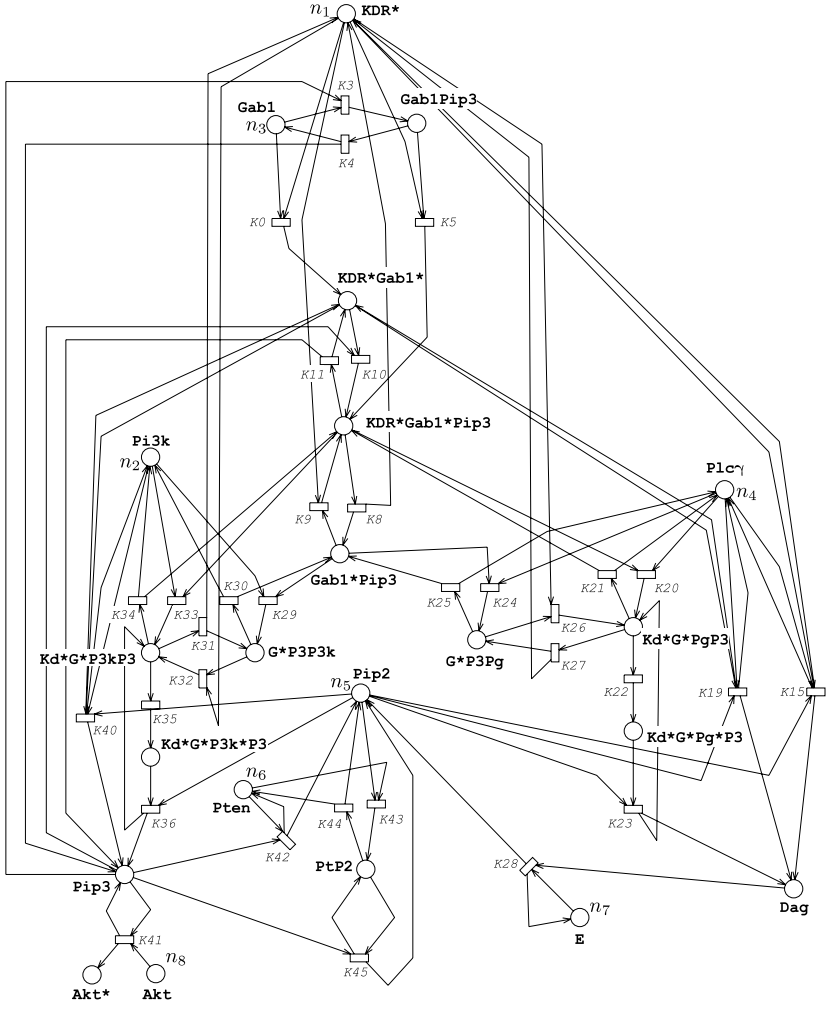

Fig. 7. SPN obtained by the second process of simplification.

5. Structural validation

Recalling that the structural properties we are interested in are P- and T-semiflows, we first verify that none of those important for the analysis of the net are lost due to the simplification. Considering the P-invariants first, every P-semiflow of the original detailed model has a counterpart in the set of P-semiflows of the simplified models (see Appendix). Indeed,

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

Table 1
Space state size of SPN models as function of the number of tokens in places KDR^*, $Gab1$, Plc_γ, E, Pip_2, $Pten$, $Pi3k$ and Akt (All these places have the same number of tokens in the initial marking).

Number of tokens	Space state size		
	Original	First simplification	Second simplification
1	110	73	27
2	6587	2601	402
3	188378	44852	3370
4	3345587	496425	19965
5	42734935	4055250	93513
6	426590775	26418100	368410
7	3505084137	144086000	1268724

during the simplification process we verified that any RC and SC has a number of P-semiflows identical to that of the corresponding reduced block.[2] Moreover, both the original and the simplified SPNs are covered by P-semiflows meaning that they are bounded. Obviously, if we consider the P-semiflows of the original model and their counterparts in the simplified ones, we find that the supports of the P-semiflows coming from the simplified models are always included in those of the original one.
These results provided by the P-semiflow analysis, show that the substitutions performed during both the simplification processes have no impact on the conservation laws, supporting the adequacy of the reduction step.

As the quantities are preserved over a P-semiflow, the same quantity will be distributed over a larger number of places in the original model than in the simplified ones. These peculiarities must be accounted for when choosing initial quantities for the simplified models and when comparing performance indexes computed for the original and simplified models, as we will briefly discuss in Section 7.3.

Considering now the T-semiflows, we can observe that not all the transitions of the original model are covered by T-semiflows. In particular, this happens for transitions K_2, K_7, and K_{30}. This identifies two different features of our model:

1. Transitions K_7 and K_2 represent the production of $KDR^*{:}Gab1^*{:}Pip_3$ and $KDR^*{:}Gab1^*$, respectively. They correspond to the initial actions of the signal transduction cascade that lead to a stationary behavior of the model which involves the proliferation and survival pathways.
2. Transition K_{30} is included only for experimental purposes to capture the activation of Akt. As mentioned in Section 3.2, the activation of Akt is completely governed by Pip_3 that is thus the quantity relevant for our studies.

Obviously, the T-semiflows involving only transitions that have been removed in the simplified models disappear. For instance, the detailed SPN has the T-semiflow $\{K_{19}, K_{20}\}$ which is not present in the SPN when sub-net RC3 (Fig. 3) including K_{19}, K_{20} and K_{21} is reduced. The T-semiflows including both removed and non-removed transitions are maintained with the removed transitions replaced by the corresponding new (equivalent) transitions. For instance, the detailed SPN has the T-semiflow $\{K_{18}, K_{19}, K_{21}, K_{58}, K_{60}\}$ which after the reduction of the sub-net RC2 with K_{19}, K_{20}, K_{21} and of the sub-net RC3 with K_{58}, K_{59}, K_{60} is still present in the form $\{K_{18}, K_{57}, K_{61}\}$ where K_{57} and K_{61} are the transitions resulting from the substitution of the corresponding RCs.

The consistency among the invariant properties of all the nets shows that these simplification processes preserve the important structural properties of the model. Moreover, the substitutions applied to the original SPN model provide new representations whose analysis has a reduced computational complexity.

In Table 1, we report the size of the state spaces of the original and of the two simplified SPNs as functions of the initial markings which have been chosen for illustrative purposes, only to show the impact that the simplifications may have on the sizes of the state spaces of the models.

For the analysis of the state space we used a version of the GreatSPN tool [26] enhanced through the use of Binary Decision Diagram data structures provided by an existing open-source library (Meddly) [24]. With this new version of GreatSPN we are able to compute the reachability set of all the models (original, first simplification, and second simplification) when in the initial marking places KDR^*, $Gab1$, Plc_γ, E, Pip_2, $Pten$, $Pi3k$ and Akt all have from 1 to 7 tokens each.

6. Quantitative analysis

The nets obtained by the reduction of both RC and SC sub-nets contain single transitions that play the role of sets of transitions of the detailed model. In order to use the simplified models, the parameters of these new transitions have to be defined. Our objective is to devise a method that allows the computation of the parameters of these substitutive transitions starting from the firing rates of all the transitions of the corresponding RC and SC sub-models in such a way that the quantitative behaviors of the simplified models remain similar to that of the detailed one.

[2] Including the test places that are often omitted in the computation of the P-semiflows.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

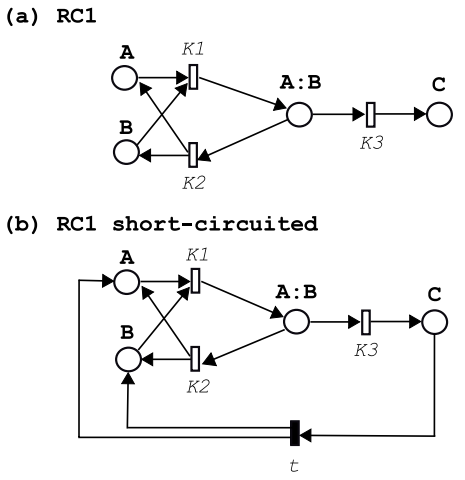

Fig. 8. The RC1 net (upper) and the short-circuited one (lower).

Since we do not have precise information about the parameters of the original detailed model (initial quantities of the different species and transition speeds), we must use a parametrization technique that is robust with respect to a large range of parameter values. The use of the CTMC that underlies the whole model as a reference for the assessment of the quality of the simplifications is unfeasible for more than 4 tokens in each of the basic places considered in the initial marking due to its state space explosion, as reported in Table 1. We thus propose to validate our technique based on the concept of *flow equivalent server* by solving the ODE derived from the different SPN models.

6.1. Flow equivalent server

The method that we use to compute the parameters is based on the concept of the flow equivalent server that was originally devised for the analysis of electrical circuits (known in that field as Norton's theorem) and that was subsequently adapted to the study of queuing networks [7,10,36]. The idea behind this concept is to consider each sub-model individually (and in isolation), to evaluate the flow of tokens that move from the input to the output places of the sub-model, and to derive information from this flow in order to define the firing rate of the single (substitutive) transition that appears in the reduced sub-model. The objective of the substitution is to ensure that the two representations provide equivalent flows of tokens from their input to output places, when operating within identical external conditions (or environments). The substitution results in (exact) equivalent behavior for the overall model only under specific stringent conditions, which our models do not usually satisfy, and only in steady state. However, as we will show in Section 7, the resulting simplified model provides a good approximation of the behavior of the detailed one in many cases.

The speed of the equivalent transition is assumed to be the throughput of tokens reaching the output place, given a specific marking of the input places of the same sub-model. In this way the equivalent firing rate assumes different values for each possible initial marking of the input places and is computed assuming a stationary behavior that is guaranteed by making sure that, when the sub-model is analyzed in isolation (i.e., disconnected from the rest of the model), each token that reaches the output place is immediately replaced by new ones in the input places. This is operationally obtained by short-circuiting the output and input places and by computing the flow along the short-circuit. In our case, the short-circuit is obtained introducing an *immediate* transition [1] which connects the output to the input places.

In order to explain this approach in detail, let us consider sub-net RC1 (see Fig. 8(a) which is identical to Fig. 4(a) and is repeated here only for convenience), where places and transitions have been renamed with general denotation. We can describe the role of this sub-net as follows. When embedded into an "overall" model (e.g., the model of Fig. 3), this component receives in input, tokens in places A and B from other parts of the whole model and moves these tokens toward C. The substitutive transition K_n (see Fig. 4(d)) performs the same action and our objective is that of associating with this transition a parameter that allows for this flow of tokens to be identical to that of sub-net RC1.

It is obvious that the parameter of K_n depends on the original parameters of RC1, k_1, k_2 and k_3. Moreover, since the more tokens arrive to places A and B the more tokens are moved to place C, the parameter of K_n has to depend also on the actual number of tokens present in places A and B.

Fig. 8 shows how this method is applied to our archetype sub-model and how the Generalized SPN [1] is constructed in order to analyze the sub-model in isolation. As we observed before, the two models of Fig. 8(a) and (b) have identical P-semiflows. For what concerns the T-semiflows, the addition of the short-circuit has the effect of making the model of Fig. 8(b) covered by T-invariants thus making the initial marking a "home state" which is a necessary condition for its steady state analysis.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

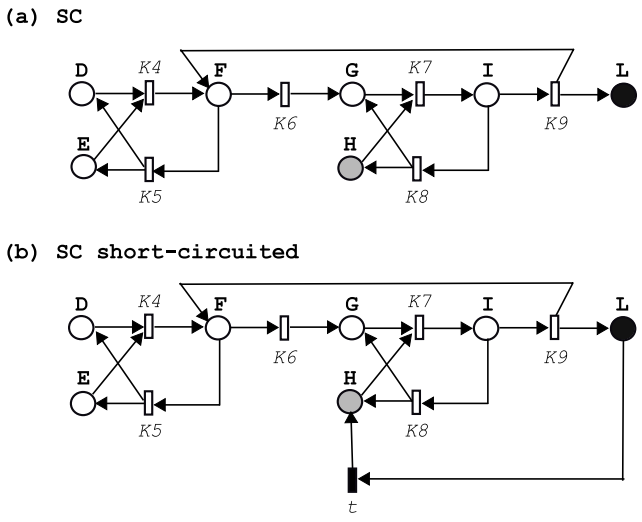

Fig. 9. Sub-net SC short-circuited.

More interesting is the representation of sub-net *SC* that is depicted in Fig. 9. Again the two sub-models have identical *P*-semiflows. Focusing the attention on the third *P*-semiflow that includes places H, I, and L, we notice that the short-circuit must connect places L and H yielding a representation (Fig. 9(b)) that is now covered also by *T-invariants*.

For this model, the flow equivalent server characterization amounts thus to the computation of the throughput of transition t of Fig. 9(b) for different and increasing configurations of the initial markings of the *input* and *enabling* places. These different transition speeds are accounted for in the SPN representing the simplified model by making the rate of the substitution transition marking dependent (i.e., a different transition speeds for every possible enabling marking).

The concept of flow equivalent server applied when the model is studied with the deterministic approach can result in symbolic expressions for the parametrization. Having at hand a symbolic expression for the parameter of the equivalent transition allows for simple analysis of the characteristics of the sub-net (for example, one can derive the maximal speed with which a given entity is produced by the sub-net). Moreover, the symbolic expression can be used to incorporate further biological assumptions into the model (one can study, for example, the effect of having in the system an overwhelming amount of a given enzyme). These possibilities can lead to new insights into the kinetics of the considered models.

In the following subsection, we illustrate the application of the flow equivalent server method by considering the sub-net RC1 because it results in quite compact analytical expressions. In Section 6.3, we will show that the sub-net SC can be treated as well with the same method, yielding however much more complex analytical expressions.

6.2. Flow equivalent server applied to ODE

As mentioned above, in order to define the parametrization of the substitutive transition K_n, we analyze the short-circuited version of RC1, depicted in Fig. 8. In particular, we derive the rate of K_n in the presence of specific numbers of tokens in places A and B, which we will denote by M_A and M_B, respectively. In line with the usual application of the method of the flow equivalent server, this is done by calculating the steady state throughput of the immediate transition t. This throughput, which depends on M_A, M_B, k_1, k_2 and k_3, will then be assigned to transition K_n in a marking dependent manner. We assume that the net is started with initial condition $X_A(0) = M_A, X_B(0) = M_B, X_{A:B}(0) = 0$, and $X_C(0) = 0$. We will denote the steady state measures of the places by X_A, X_B, $X_{A:B}$ and X_C. The ODEs describing the evolution of the short-circuited RC1 have to reflect the effect of the immediate transition t. The effect is twofold: first, since tokens are immediately moved by t, the amount of tokens in place C is constantly zero; second, transition t puts tokens into places A and B which can be seen as if transition K_3 was moving tokens from place A:B directly to places A and B. Consequently, the steady state measure of place C is zero (i.e., $X_C = 0$) while the steady state measure of the other places can be determined by considering

- the fact that in steady state the rate of change of the quantities contained in the different places is zero, i.e., we have

$$\begin{aligned} \frac{dX_A(t)}{dt} &= 0 = -k_1X_AX_B + k_2X_{A:B} + k_3X_{A:B}, \\ \frac{dX_B(t)}{dt} &= 0 = -k_1X_AX_B + k_2X_{A:B} + k_3X_{A:B}, \\ \frac{dX_{A:B}(t)}{dt} &= 0 = +k_1X_AX_B - k_2X_{A:B} - k_3X_{A:B} \end{aligned} \tag{2}$$

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

where the term $k_3X_{A:B}$ in the first two equations are due to the fact that tokens moved by K_3 from place $A{:}B$ arrive to places A and B immediately;

- the relations given by the invariants of the net which are

$$X_A + X_{A:B} = M_A, \tag{3}$$
$$X_B + X_{A:B} = M_B. \tag{4}$$

Eqs. (2), (3) and (4) are three independent equations for the three unknowns. There are two solutions, but only one of them guarantees positivity of the unknowns. The throughput of the immediate transition t equals the throughput of transition K_3 which is $k_3X_{A:B}$ and that turns out to be

$$\frac{k_3\left(k_1(M_A+M_B)+k_3+k_2-\sqrt{(k_3+k_2+k_1(M_A+M_B))^2-4k_1^2M_AM_B}\right)}{2k_1}. \tag{5}$$

This last result provides a symbolic expression for the parametrization of the flow equivalent transition K_n. Note that the application of (5) introduces more complicated dependences among the actual reaction rates and the actual quantities of the entities than we have in the case of the detailed model where we apply the law of the generalized mass action. Indeed, the ODEs describing the simplified models are no longer of the form given in (1).
However, it is worth to note that by introducing

$$k_M = \frac{k_2+k_3}{k_1},$$

coinciding with the constant of Henri–Michaelis–Menten [35], Eq. (5) can be rewritten as

$$\frac{k_3\left(M_A+M_B+k_M-\sqrt{(M_A-M_B)^2+2k_M(M_A+M_B)+k_M^2}\right)}{2}, \tag{6}$$

which contains one parameter less. This means that every application of the simplification step to a sub-net of type RC1, the number of parameters required for the analysis of the system decreases by one. Another point to be observed is that as M_A (or M_B) tends to infinity, the speed of the equivalent server tends to k_3M_B (to k_3M_A). These limits provide the maximal speed of the substitutive reaction and they coincide with the maximal speed obtained by applying Henri–Michaelis–Menten kinetics regarding enzymatic reactions. Note that the way we derived the maximal speed of the RC1 sub-net did not require any biological assumptions apart of those expressed by the invariants given in (3) and (4).

The approximate kinetic, which we derived above based on the concept of the flow equivalent server, coincides with the approximation introduced in [38]. The novelty here is, however, twofold. First, our derivation comes from an automatizable procedure that can be applied to other type of sub-nets as well (as we will show for the SC sub-net in Section 6.3). Second, we employed this approximation in a systematic manner within a complex networks to every (structurally) identical sub-net.

Let us mention here that in [2] we investigated in detail the flow equivalence based approximate kinetics and showed that it gives satisfactory approximation not only in the standard deterministic setting, but also in the case when the behavior is modeled by a stochastic process.

6.3. Flow equivalent server for sub-net SC

The sub-net SC can be handled in the same way as the sub-net RC1. In this case, there are six transitions (K_4, K_5, K_6, K_7, K_8 and K_9), one input place (H), and two enabling places (D and E) (see Fig. 9). Consequently, the parametrization of the substitutive transition takes into account the actual amount of tokens in these three places (denoted by M_H, M_D and M_E) and depends on the rates of these six transitions (denoted by k_4, k_5, k_6, k_7, k_8 and k_9). The parametrization according to flow equivalence leads again to a symbolic expression which, even if much more complicated than in the case of the sub-net RC, can be obtained, applied and manipulated with the help of tools for symbolic computations, like Maple and Mathematica.

In Fig. 10, we report the numerical values of the rate of the flow equivalent server for some sets of parameters. Looking at these diagrams it can be seen that having a limited number of tokens in just one of the involved places (it does not matter which) limits the maximum speed even if the number of tokens in the other places increases.

7. Validation and results

The simplification process proposed in Section 4.1 results in SPNs which maintain certain qualitative properties (i.e. invariants) of the original SPN, but are approximations of the detailed model from a quantitative point of view. In this section, we report on several modeling experiments that were performed in order to check the validity of the simplifications from the point of view of quantitative measures. Indeed, before using the simplified models in a parameter identification experiment which uses real data coming from wet-lab experiments, it is necessary to make sure that a satisfactory correspondence exists

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

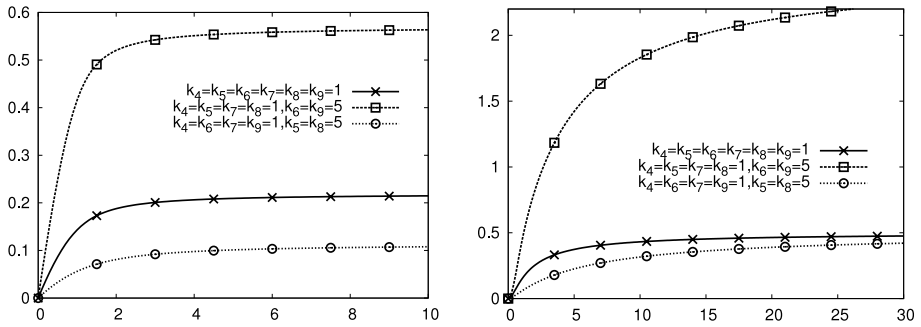

Fig. 10. Sub-net SC: speed of the flow equivalent server as function of M_D with $M_E = M_H = 1$ on the left and as function of $M_D = M_H$ with $M_E = 1$ on the right and different values of k_4, k_5, k_6, k_7, k_8, k_9.

among the quantitative temporal behaviors of the detailed and the simplified models for a wide range of (possible) model parameters. This test allows the development of confidence about the fact that the reduced model is suited for a preliminary analysis of the angiogenic signal transduction events through in silico experiments.

The accuracy assessment was performed applying the deterministic approach described in Section 2.3. For several different sets of parameters, we compared the temporal behaviors of the detailed SPN with those of the simplified SPNs. Throughout the comparisons we concentrated on three important entities: Pip_3 and *Dag* which are, in our model, the most important indicators of survival and proliferation, respectively, and Pip_2 which is the common substrate to both the survival and the proliferation pathways.

A preliminary step of assessment had already been reported in the original version of this paper [29]. As the idea of connecting the parameters of the detailed model to the parameters of the simplified ones by the concept of equivalent server had not yet been developed, at that time, we performed the assessment either leaving the speed of the reactions unchanged or determining them by numerical optimization that required the analysis of the whole model. The introduction of the concept of equivalent server allows for parameterizing the simplified models without numerical optimization. Moreover, this way of parametrization is modular in the sense that in order to parameterize a given substitutive transition only the corresponding sub-net has to be analyzed, not the whole model.

7.1. First experimental set

In this section, we report on the results concerning the two cases that were already used in [29] and then in Section 7.2 we illustrate the robustness of the approach on cases where the parameters are chosen in a random manner. For all of the presented examples, the initial situation is the same: for all three SPNs (see Figs. 3, 5 and 7), we use the initial marking $n_1 = 2$, $n_2 = n_3 = n_4 = 1$, $n_5 = 20$, $n_6 = n_7 = n_8 = 1$, (which indicate the initial number of tokens in places *KDR**, *Pi*3*k*, *Gab*1, Plc_γ, Pip_2, *Pten*, *E*, and *Akt*, respectively), which reflects the amount differences that are likely to exist in wet-lab experiments.

For the first two cases, two different sets of transition rates are used to push the behaviors of the models in opposite directions. In the first case, the rates are such that the transitions along the survival pathway are ten times faster than all the others. Fig. 11 depicts the temporal behavior of the entities Pip_3, *Dag* and Pip_2 for both the detailed model and the simplified ones. With these parameters the amount of Pip_3 increases, the amount of Pip_2 decreases and the amount of *Dag* initially increases, but remains relatively low. For all the three quantities, the behaviors in the three models are similar in the sense that the entities move along curves of the same shape and the changes take place in the same time intervals. The final amounts are different for the three models, but this is not surprising as the number of places is different in the three models and this implies that, considering the mass of each individual component appearing in the models, in steady state it is distributed over (subdivided among) different numbers of locations in the different models (as already mentioned, we will return to this point in Section 7.3).

In the second case, the transitions along the proliferation pathway are ten times faster than all the others. Fig. 12 reports the temporal behavior of the three models for the three selected quantities. In this case as well, the shapes of the curves show a good correspondence. For what concerns instead the time intervals in which the changes occur, the more simplified model is somewhat "faster" than the detailed one.

7.2. Randomly generated parameters

In the rest of this section, we report on examples with random reaction rates. The reaction rates of the detailed model are chosen according to the formula

$$(1 + 9r_1) \times 10^{-a+2ar_2} \tag{7}$$

where r_1 and r_2 are random numbers with uniform distribution on [0, 1]. Accordingly, $(1 + 9r_1)$ falls in the interval [1, 10] while $-a + 2ar_2$ in the interval $[-a, +a]$. Consequently, for a given value of a, the reaction rates cover at most $2a + 1$ orders

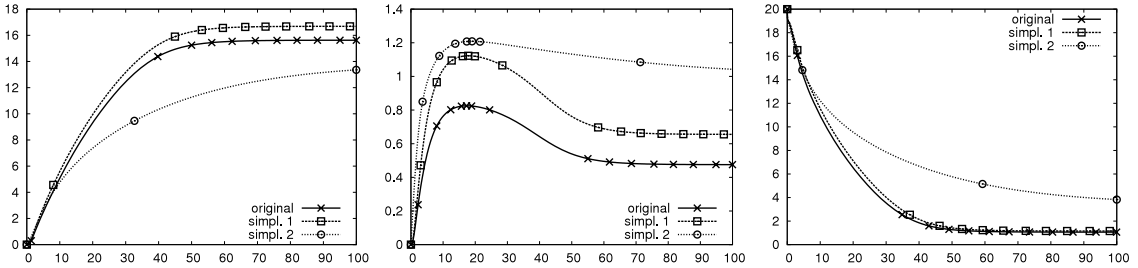

Fig. 11. Behavior of Pip_3, Dag and Pip_2 with the first set of parameters in the detailed and in the two simplified models.

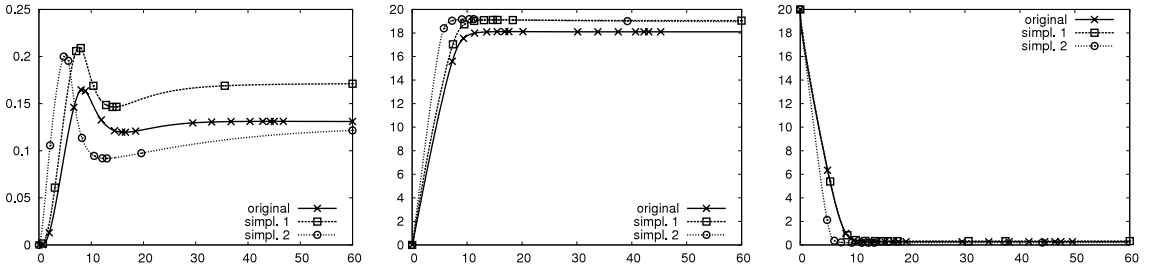

Fig. 12. Behavior of Pip_3, Dag and Pip_2 with the second set of parameters in the detailed and in the two simplified models.

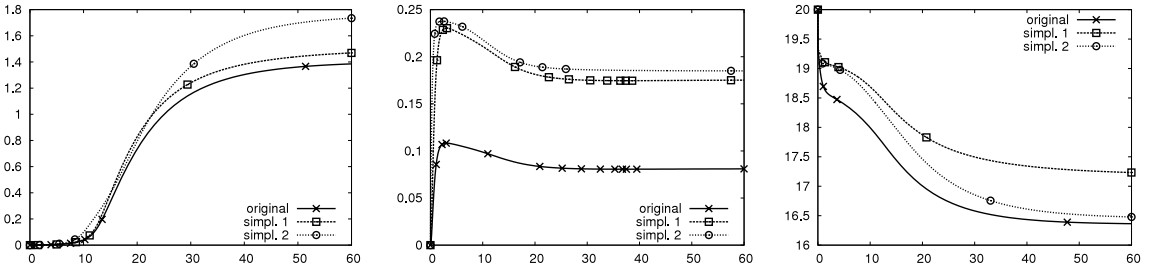

Fig. 13. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 1 and $a = 1$ in (7).

of magnitude. In general, a higher value of a implies more stiffness of the differential equations and more complex behavior of the model. Moreover, since it is shown in [38] that the approximate kinetic is appropriate on a single RC type sub-net if $k_3/(k_2 + k_3)$ is small, we require for every RC type sub-net $k_3/(k_2 + k_3) < 0.2$ (meaning that k_2 is at least four times larger than k_3) and we generate the random rates of all the transitions of the detailed model randomly, but with the restriction of the previous inequality being satisfied for all the transitions of the RC blocks recognized in the net.

Note that the above manner of parameterizing the detailed model (see Eq. (7)), introduces extreme irregularities and puts the simplification process under stress conditions that would hardly occur with the reaction parameters of realistic cases. Moreover, to validate the robustness of the approach in a general setting, we ran several sets of experiments using different seeds for the random number generator. Figs. 13–18 are representative of the results that we obtained with these experiments and depict the dynamics observed for two different values of the seed and for $a = 1, 2, 3$. With the first seed the effect of a can clearly be seen: with $a = 1$ the three models behave very similarly; with $a = 3$, the major characteristics of the curves are maintained, but the shapes of the curves differ more. In this case, there is not much difference between the two simplified models. With the second seed the difference in robustness between the two simplified models can be observed. The more simplified model exhibits quite different behavior for what concerns Pip_3 with $a = 2$ and $a = 3$. The less simplified model provides a good approximation even if it is somewhat "slower" than the detailed model.

7.3. Correction of absolute amount in the simplified models

From Figs. 11–18 one can observe that, even if the major characteristics of the original model are maintained by the simplified ones, the absolute amount of the entities can be quite different. This is natural as in the simplified models the same original quantity of substance is distributed among less places. This problem can be alleviated by a post-correction. The complete treatment of this correction approach is beyond the scope of this paper and will be tackled in future works, but we aim to illustrate it in a simple case in order to show that the proper interpretation of the results can lead to a substantially improved match between the simplified models and the original one.

Let us consider the RC sub-net depicted in Fig. 4. In its simplified version the place in the middle is eliminated. The substance that is present in this place in the original model is blocked in the input places of the RC sub-net in the simplified

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

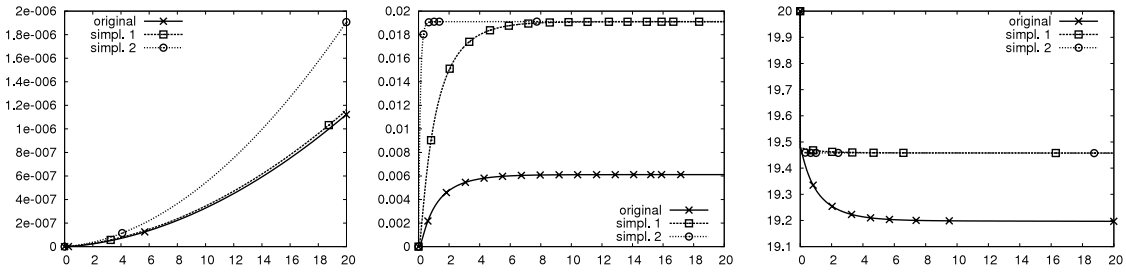

Fig. 14. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 1 and $a = 2$ in (7).

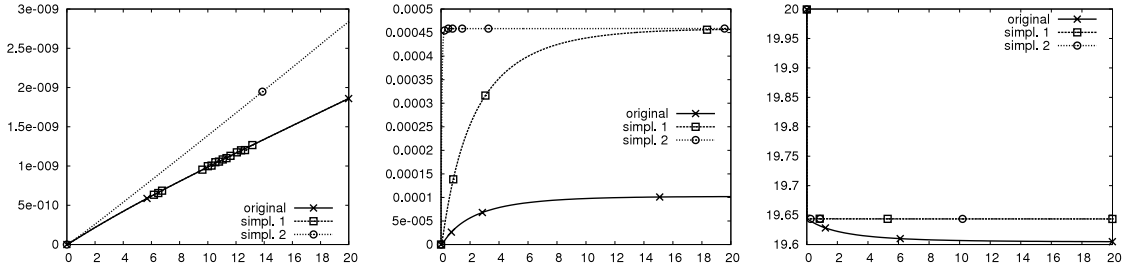

Fig. 15. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 1 and $a = 3$ in (7).

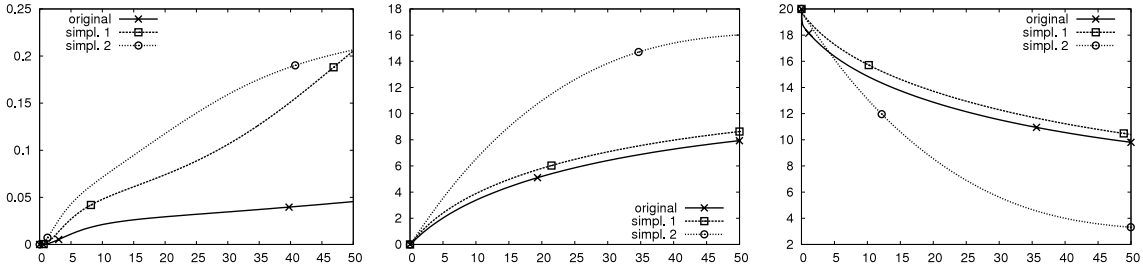

Fig. 16. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 2 and $a = 1$ in (7).

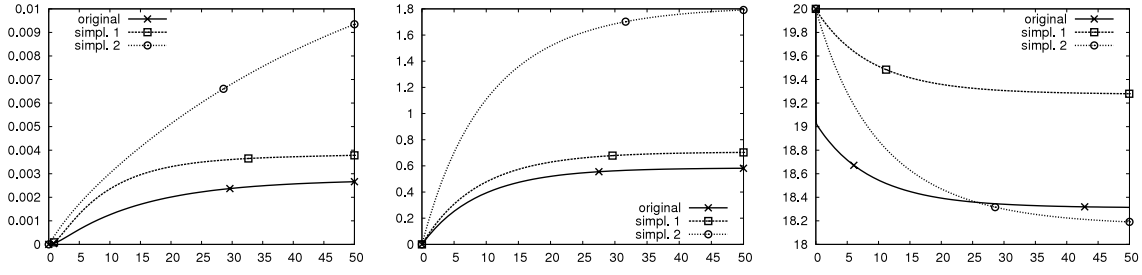

Fig. 17. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 2 and $a = 2$ in (7).

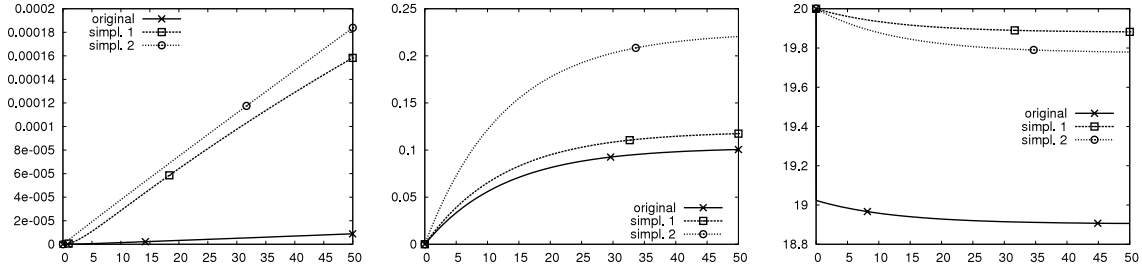

Fig. 18. Behavior of Pip_3, Dag and Pip_2 with random rates with seed equals 2 and $a = 3$ in (7).

model. Accordingly, there is more substance present in these input places in the simplified model than in the original one. The quantity of substance that should be in the eliminated place can be approximated by the quantity of substance present in place $A{:}B$ in steady state in the short-circuited version of RC (depicted in Fig. 8). This is the amount, given by the

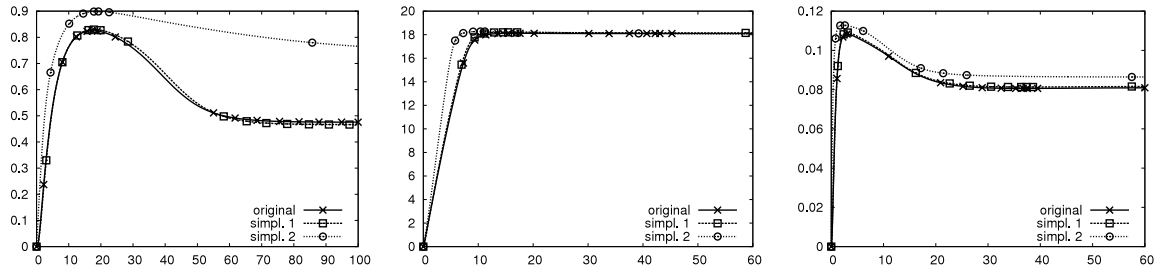

Fig. 19. Corrected behavior of *Dag* with the first and the second set of parameters (left and middle, respectively) and with random rates with seed equals 1 and $a = 1$ (right).

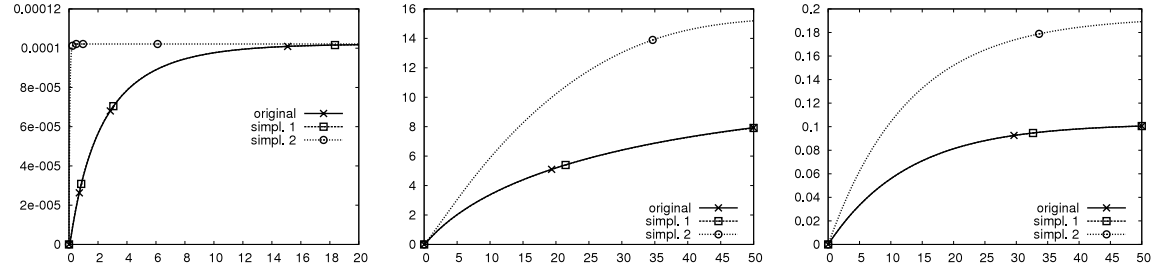

Fig. 20. Corrected behavior of *Dag* with random rates with seed equals 1 and $a = 3$ (left) and with seed equals 2 and $a = 1, 3$ (middle and right, respectively).

expression in (6) divided by k_3, by which the quantities present in the input places have to be decreased. Naturally, this correction procedure is consistent for those places which play the role of input place in only one sub-net. One such place is *Dag* positioned in the bottom right part of Fig. 3.

We applied the correction to the experiments with random rates (Figs. 13–18) because for these the mismatch between the original and the simplified models is larger than that in case of the experiments with structured rates (Fig. 11–12).

As one can observe in Figs. 19 and 20, the correction leads to a very precise match among the results predicted by the less simplified model and the original one in all cases that we have tried and thus even in the cases of parameter choices that put the model in extreme situations. The correction method is not so effective in the cases of the results obtained with the more simplified model. In this case, the post-correction improves the quality of the results, but is not able to always provide the surprisingly good match mentioned before. Additional work is needed to obtain satisfactory (and robust) results even in this case, but we believe that more sophisticated post-correction techniques can be developed in order to have a good match with the more simplified model as well.

8. Conclusions and future work

The main goal of this paper is to propose a robust way of simplifying models that describe complex biological systems and to define a method that can lead to a manageable model. The simplification process exploits the concept of flow equivalent server used to compute the parameters of the simplified model.

As mathematical models of metabolic pathways have to deal with both mass conservation and kinetic aspects of the modeled phenomenon, the simplification procedure as well has to take into account both of these aspects.

Since the P-invariants computed from Petri nets can be seen as the counterpart of the law of conservation of mass, we propose simplifications that maintain P-invariants and hence guarantee that the laws of conservation of mass expressed in the original model are maintained throughout all the reduction steps.

Kinetics describe the dynamics of the model. In the case of metabolic pathways, the most commonly applied kinetics is provided by the law of generalized mass action. The parametrization method that we have used to define the characteristics of the simplified models, is based on the concept of flow equivalence, and is thus aimed at mimicking the kinetics of the original model. The application of flow equivalence can be seen as analyzing the flux balance of the original model. In this respect, it is similar to evaluate the physicochemical constraints of the metabolic network under study.

A clear advantage of the proposed approach is that it allows for parameterizing the simplified model block by block instead of analyzing the original model as a whole. Moreover, the obtained parameterizations are general in the sense that they can be used in any other model that contains components with net structures similar to those discussed in this paper.

A promising result is that the application of flow equivalence leads to a symbolic expression that describes the throughput of the considered sub-net. These symbolic expressions provide a simple tool to introduce biological assumptions into the analysis of the model which can provide new insights into the kinetics underlying the studied phenomena. In particular, by analyzing the expression of the equivalent flow of the RC sub-net, we have provided a connection between the behavior of the RC sub-net and the kinetics of Henri–Michaelis–Menten applied to enzymatic reactions. In the future, we plan to work

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

Table A.2
P-semiflows of the **original** SPN model.

Semiflow ID	Semiflow support
1	$Kd^*G^*Pi3k - Kd^*G^*Pi3k^* - Kd^*G^*Pi3k^*P2$ $-Kd^*G^*Pi3k^*P3P2 - Kd^*G^*Pi3k^*P3 - Kd^*G^*Pi3kP3$ $-Kd^*Pg^*P2 - Kd^*Pg^* - Kd^*Pg - Kd^*G^*Pg$ $-Kd^*G^*Pg^* - Kd^*G^*Pg^*P2 - Kd^*G^*PgP3$ $-Kd^*G^*Pg^*P3 - Kd^*G^*Pg^*P3P2 - KDR^*Gab1^*Pip_3$ $-KDR^*Gab1^* - KDR^*Gab1Pip_3 - KDR^*Gab1 - KDR^*$
2	$Kd^*G^*Pi3k - Kd^*G^*Pi3k^* - Kd^*G^*Pi3k^*P2$ $-Kd^*G^*Pi3k^*P3P2 - Kd^*G^*Pi3k^*P3 - Kd^*G^*Pi3kP3$ $-G^*Pi3kP3 - Kd^*G^*Pg - Kd^*G^*Pg^* - Kd^*G^*Pg^*P2$ $-G^*PgP3 - Kd^*G^*PgP3 - Kd^*G^*Pg^*P3 - Kd^*G^*Pg^*P3P2$ $-Gab1^*Pip_3 - KDR^*Gab1^*Pip_3 - KDR^*Gab1^* - KDR^*Gab1Pip_3$ $-KDR^*Gab1 - Gab1Pip_3 - Gab1$
3	$Kd^*G^*Pi3k - Kd^*G^*Pi3k^* - Kd^*G^*Pi3k^*P2$ $-Kd^*G^*Pi3k^*P3P2 - Kd^*G^*Pi3k^*P3 - Kd^*G^*Pi3kP3$ $-G^*Pi3kP3 - Pi3k$
4	$Kd^*Pg^*P2 - Kd^*Pg^* - Kd^*Pg - Kd^*G^*Pg$ $-Kd^*G^*Pg^* - Kd^*G^*Pg^*P2 - G^*PgP3 - Kd^*G^*PgP3$ $-Kd^*G^*Pg^*P3 - Kd^*G^*Pg^*P3P2 - Plc_\gamma$
5	$-Kd^*G^*Pi3k^*P2 - Kd^*G^*Pi3k^*P3P2 - Kd^*G^*Pi3k^*P3$ $-Kd^*G^*Pi3kP3 - G^*Pi3kP3 - Kd^*Pg^*P2 - Kd^*G^*Pg^*P2$ $-G^*PgP3 - Kd^*G^*PgP3 - Kd^*G^*Pg^*P3 - Kd^*G^*Pg^*P3P2$ $-Gab1^*Pip_3 - KDR^*Gab1^*Pip_3 - KDR^*Gab1Pip_3 - Gab1Pip_3 - Pip_3$ $-AktP3 - Dag - DagE - Pip_2 - PtP3 - PtP2 - PtP3P2$
6	$PtP3 - Pten - PtP2 - PtP3P2$
7	$Akt - Akt^* - AktP3$
8	$DagE - E$

in this direction by analyzing further the symbolic expressions obtained for the two sub-nets considered in this paper in order also to derive other symbolic expressions for different types of sub-nets.

Future work must include a more detailed analysis of the accuracy of applying the concept of flow equivalent server in the parametrization of the simplified models. The methodology of substituting reactions with a single one by flow equivalence is based on the idea of isolating the sub-net we want to substitute from the rest of the model and of short-circuiting its output and input places to study its behavior in an equilibrium environment. The analysis provides the steady state features of the sub-net and these features are then applied to parameterize the substitutive transition. As the substitution is based on steady state measures, the correspondence it guarantees between the steady state behavior of the detailed model and the steady state behavior of the simplified model is stronger than that between the transient behaviors of the two models. Since in systems biology the transient characteristics are often of crucial interest, it is important to check whether flow equivalence based on steady state is sufficient to provide robust simplification in a wide range of cases.

In this work, we have shown that the accuracy of the method is satisfactory within a very large set of model parameters that we generated randomly to stress the robustness of the test of our approach. When we applied the simplification process on the RC sub-nets only, yielding what we called the *simplified* 1 model, the dynamics of the original model was captured in a very good manner. Moreover, initial results are provided to show that correcting actions can be performed to improve the quality of the results considering the way in which the substance may distribute over the smaller number of places represented in the simplified models. The methodology devised for performing these correcting actions accounts for the structure of the simplified sub-model, but is independent of the values of its parameters. When applied in this case the correction yielded surprisingly accurate results. For what concerns the *simplified* 2 model the trend of the evolution of the species amount is captured in an appropriate manner, but the accuracy is much less satisfactory. In this case the correcting actions mentioned before improve the quality of the results, but the accuracy remains questionable and future work is needed to improve the correction methodology when applied to sub-models with a more complex structure as is the case for those involved in the construction of this more simplified model.

Other ways of simplifying the detailed model can also be studied in the future. One direction is suggested by the reachability graph of the SPN representing the detailed model in which the majority of the states are transient. For what concerns steady state analysis, this suggests to identify the subcomponents that are alive in stationary conditions and use this information in the simplification process.

Even if the simplification procedure is performed without taking into explicit account assumptions coming from biological knowledge of the phenomenon under study, it is worthwhile to investigate if it can be interpreted from a biological

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013

Table A.3
P-semiflows of the SPN model obtained after the **first** simplification process.

Semiflow ID	Semiflow support
1	$Kd^*G^*P3Pi3k^* - Kd^*G^*Pi3k^* - Kd^*G^*P3Pg^*$ $-Kd^*G^*Pg^* - Kd^*Pg^* - KDR^* - KDR^*G^* - KDR^*G^*P3$ $-Kd^*G^*Pi3k - Kd^*G^*Pg - Kd^*Pg - Kd^*G^*P3Pi3k - Kd^*G^*P3Pg$
2	$G^*P3Pi3k - Kd^*G^*P3Pi3k^* - Kd^*G^*P3Pi3k - Kd^*G^*Pi3k^* - Kd^*G^*Pi3k - G^*P3Pg$ $-Kd^*G^*P3Pg^* - Kd^*G^*P3Pg - Kd^*G^*Pg^* - Kd^*G^*Pg - G^*P3 - Gab1 - GP3$ $-KDR^*G^* - KDR^*G^*P3$
3	$G^*P3Pi3k - Kd^*G^*P3Pi3k^* - Kd^*G^*Pi3k^* - Pi3k - Kd^*G^*P3Pi3k - Kd^*G^*Pi3k$
4	$G^*P3Pg - Kd^*G^*P3Pg^* - Kd^*G^*Pg^* - Kd^*Pg^* - Plc_\gamma$ $-Kd^*G^*Pg - Kd^*Pg - Kd^*G^*P3Pg$
5	$G^*P3Pi3k - Kd^*G^*P3Pi3k^* - G^*P3Pg - Kd^*G^*P3Pg^*$ $-Dag - Pip_2 - G^*P3 - Pip_3 - GP3 - KDR^*G^*P3 - PtP2 - Kd^*G^*P3Pi3k - Kd^*G^*P3Pg$
6	$Pten - PtP2$
7	$Akt - Akt^*$
8	E

Table A.4
P-semiflows of the SPN model obtained after the **second** simplification process.

Semiflow ID	Semiflow support
1	$Kd^*G^*Pi3k^*P3 - Kd^*G^*Pg^*P3 - KDR^*G^*P3 - KDR^*G^*$ $-KDR^*Kd^*G^*Pi3kP3 - Kd^*G^*PgP3$
2	$Kd^*G^*Pi3k^*P3 - G^*Pi3kP3 - G^*PgP3 - Kd^*G^*Pg^*P3 - KDR^*Gab1^*$ $-KDR^*G^*P3 - GP3 - G^*P3 - Gab1 - KDR^*G^*Pi3kP3 - KDR^*G^*PgP3$
3	$Kd^*G^*Pi3k^*P3 - G^*Pi3kP3 - Pi3k - Kd^*G^*Pi3kP3$
4	$Kd^*G^*Pg^*P3 - G^*PgP3 - Plc_\gamma - Kd^*G^*PgP3$
5	$G^*Pi3kP3 - G^*PgP3 - Kd^*G^*Pg^*P3 - Kd^*G^*Pi3k^*P3 - KDR^*G^*P3$ $-GP3 - G^*P3 - Pip_3 - Dag - Pip_2 - PtP2 - Kd^*G^*PgP3 - Kd^*G^*Pi3kP3$
6	$Pten - PtP2$
7	$Akt - Akt^*$
8	E

point of view as well. For instance, the substitution applied to RC1 can be translated into the following biological assumption: the docking of a signal protein (such as *Gab*1) to the proper upstream activator (such as *KDR**) is simultaneous to the activation of the signal protein. Moreover, the simplification of SCs can be biologically interpreted as if the formation of the product (such as *Dag*) was independent of the reagent's identity. For this reason, we plan to work more on this simplification process and to improve it by considering not only the description of the model, but the related biological assumptions as well.

Acknowledgements

This work was supported by grants from Italian Association for Cancer Research, the Italian Ministero dell'Università e della Ricerca, the University of Torino, EUCAAD 200755. The authors are grateful to the anonymous referees for the many helpful suggestions and comments that have led to important improvements in the presentation of this paper. Furthermore, the authors wish to thank Steve Bruell for the contribution that he gave in the definition of the model during the very early stage of this work.

Appendix. *P*-Semiflows

See Tables A.2–A.4.

References

[1] M. Ajmone Marsan, G. Balbo, G. Conte, A class of generalized stochastic Petri nets for the performance analysis of multiprocessor systems, ACM Transactions on Computer Systems 2 (1) (1984).
[2] A. Angius, G. Balbo, F. Cordero, A. Horvath, D. Manini, Comparison of approximate kinetics for unireactant enzymes: Michaelis-Menten against the equivalent server, in: Proceedings of the International Workshop on Biological Processes and Petri Nets, BioPPN, 2010.
[3] S. Baarir, M. Beccuti, D. Cerotti, M. De Pierro, S. Donatelli, G. Franceschinis, The GreatSPN tool: recent enhancements, ACM SIGMETRICS, Performance Evaluation Review 36 (4) (2009) 4–9.
[4] G. Balbo, Introduction to stochastic Petri nets, in: E. Brinksma, H. Hermanns, J.-P. Katoen (Eds.), Lectures on Formal Mathods and Performance Analysis, in: LNCS, vol. 2090, Springer, Berlin, Germany, 2001, pp. 1–37.
[5] H. Busch, W. Sandmann, V. Wolf, A numerical aggregation algorithm for the enzyme-catalyzed substrate conversion, in: Computational Methods in Systems Biology, in: LNCS, vol. 4210, Springer, 2006, pp. 298–311.
[6] M. Calder, V. Vyshemirsky, D. Gilbert, R. Orton, Analysis of signalling pathways using continuous time Markov chains, Transactions on Computational Systems Biology VI 4 (2006) 44–67.
[7] K.M. Chandy, U. Herzog, L.S. Woo, Parametric analysis of queueing networks, IBM Journal of R. & D. 19 (1) (1975) 36–42.
[8] I.-C. Chou, E. O. Voit, Recent developments in parameter estimation and structure identification of biochemical and genomic systems, Mathematical Biosciences 219 (2) (2009) 57–83.
[9] M. Dance, A. Montagner, A. Yart, B. Masri, Y. Audigier, B. Perret, J. Salles, P. Raynal, The adaptor protein Gab1 couples the stimulation of vascular endothelial growth factor receptor-2 to the activation of phosphoinositide 3-kinase, Journal of Biological Chemestry 281 (2006) 23285–23295.
[10] P.J. Denning, J.P. Buzen, The operational analysis of queueing network models, ACM Computing Surveys 10 (3) (1978) 225–261.
[11] W. Feller, An Introduction to Probability Theory and its Applications, Vol. 1, John Wiley, 1968.
[12] H. Gerber, A. McMurtrey, J. Kowalski, M. Yan, B. Keyt, V. Dixit, N. Ferrara, Vascular endothelial growth factor regulates endothelial cell survival through the phosphatidylinositol 3'-kinase/Akt signal transduction pathway. requirement for Flk-1/KDR activation, Journal of Biological Chemestry 273 (1998) 30336–30343.
[13] D. Gillespie, A rigorous derivation of the master chemical equation, Physica 188 (1992) 404–425.
[14] P. Goss, J. Pecoud, Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets, Proceedings of the National Academy of Sciences 95 (12) (1998) 6750–6755.
[15] E. Grafahrend-Belau, F. Schreiber, M. Heiner, A. Sackmann, B. Junker, S. Grunwald, A. Speer, K. Winder, I. Koch, Modularization of biochemical networks based on classification of Petri net t-invariants, BMC Bioinformatics 9 (2008).
[16] J. Heath, M. Kwiatkowska, G. Norman, D. Parker, O. Tymchyshyn, Probabilistic Model Checking of Complex Biological Pathways, Springer, 2006, pp. 32–47.
[17] M. Heiner, D. Gilbert, R. Donaldson, Petri nets for systems and synthetic biology, in: 8th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, in: LNCS, Springer, Bertinoro, Italy, 2008, pp. 215–264.
[18] M. Heiner, I. Koch, J. Will, Model validation of biological pathways using Petri nets demonstrated for apoptosis, BioSystems 75 (2004) 10–28.
[19] M. Heiner, C. Mahulea, M. Silva, On the importance of the deadlock trap property for monotonic liveness, in: Proceedings of the International Workshop on Biological Processes and Petri Nets, BioPPN, 2010.
[20] R. Hofestädt, A Petri net application of metabolic processes, Journal of System Analysis, Modeling and Simulation 16 (1994) 113–122.
[21] R. Hofestädt, S. Thelen, Quantitative modeling of biochemical networks, In Silico Biology 1 (6) (1998).
[22] T.G. Kurtz, The relationship between stochastic and deterministic models for chemical reactions, Journal of Chemical Physics 57 (7) (1972) 2976–2978.
[23] M. Laramée, C. Chabot, M. Cloutier, R. Stenne, M. Holgado-Madruga, A. Wong, I. Royal, The scaffolding adapter Gab1 mediates Vascular Endothelial Growth factor signaling and is required for endothelial cell migration and capillary formation, Journal of Biological Chemestry 282 (2007) 7758–7769.
[24] MEDDLY, 2010. Webpage. http://sourceforge.net/projects/meddly.
[25] G. Memmi, J. Vautherin, Advanced algebraic techniques, in: W. Brawer, W. Reisig, G. Rozenberg (Eds.), Advances on Petri Nets '86—Part I, in: LNCS, vol. 254, Springer Verlag, Bad Honnef, West Germany, 1987.
[26] A. Miner, J. Babar, M. Beccuti, S. Donatelli, Greatspn enhanced with decision diagram data structures, in: Proc. 31-st Int. Conf. Applications and Theory of Petri Nets ICATPN 2010, in: LNCS, vol. 6128, Springer, Braga, Portugal, 2010, pp. 308–317.
[27] M.K. Molloy, Performance analysis using stochastic Petri nets, IEEE Transaction on Computers 31 (9) (1982) 913–917.
[28] T. Murata, Petri nets: properties, analysis, and applications, Proceedings of the IEEE 77 (4) (1989) 541–580.
[29] L. Napione, D. Manini, F. Cordero, A. Horváth, A. Picco, M.D. Pierro, S. Pavan, M. Sereno, A. Veglio, F. Bussolino, G. Balbo, On the use of stochastic Petri nets in the analysis of signal transduction pathways for angiogenesis process, in: Proc. of The 7th Conference on Computational Methods in Systems Biology, CMSB 2009, in: Lecture Notes in Bioinformatics, vol. 5688, Bologna, Italy, 2009, 281–295.
[30] S. Natkin, Les réseaux de Petri stochastiques et leur application à l'évaluation des systèmes informatiques, Thèse de Docteur Ingégneur, CNAM, 1980.
[31] A. Olsson, A. Dimberg, J. Kreuger, L. Claesson-Welsh, Vegf receptor signalingin control of vascular function, Nature Reviews Molecular Cell Biology 7 (5) (2006) 359–371.
[32] V. Reddy, M. Mavrovouniotis, M. Liebman, Petri net representation in metabolic pathways, in: Proc. Int. Conf. Intelligent Systems for Molecular Biology, 1993, pp. 328–336.
[33] W. Reisig, A Primer in Petri Net Design, Springer Compass International, 1992.
[34] H. Sauro, B. Ingalls, Conservation analysis in biochemical networks: computational issues for software writers, Biophysical Chemistry 109 (1) (2004) 1–15.
[35] I. H. Segel, Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady State Enzyme Systems, Wiley, New York, 1993.
[36] W.J. Stewart, Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling, Princeton University Press, 2009 (Chapter 15).
[37] T. Takahashi, H. Ueno, M. Shibuya, VEGF activates protein kinase C-dependent, but Ras-independent Raf-MEK-MAP kinase pathway for DNA synthesis in primary endothelial cells, Oncogene 18 (1999) 2221–2230.
[38] A.R. Tzafriri, Michaelis–Menten kinetics at high enzyme concentrations, Bulletin of Mathematical Biology 65 (2003) 1111–1129.
[39] R. Vallabhajosyula, V. Chickarmane, H. Sauro, Conservation analysis of large biochemical networks, Bioinformatics 22 (2006) 346–353.
[40] R.R. Vallabhajosyula, H.M. Sauro, Complexity reduction of biochemical networks, in: Winter Simulation Conference, 2006, pp. 1690–1697.
[41] E.O. Voit, Computational Analysis of Biochemical Systems, Cambridge University Press, 2000.

Please cite this article in press as: F. Cordero, et al., Simplification of a complex signal transduction model using invariants and flow equivalent servers, Theoretical Computer Science (2011), doi:10.1016/j.tcs.2011.06.013